McCandless JW (1999) Detection of aircraft in video sequences using a predictive optical flow algorithm. *Optical Engineering*, 3: 523-530.

Detection of Aircraft in Video Sequences Using A Predictive Optical Flow Algorithm

Jeffrey W. McCandless
Human Information Processing Research Branch

Mail Stop 262-2
NASA Ames Research Center
Moffett Field, CA 94035-1000

phone: (650) 604-1162
fax: (650) 604-0255
e-mail: jmccandless@mail.arc.nasa.gov

Abstract. This paper presents a computer vision algorithm that segregates spurious optical flow artifacts to detect a moving object. The algorithm consists of six steps. First, the pixels in each image are shifted to compensate for camera rotation. Second, the images are smoothed with a spatio–temporal Gaussian filter. Third, the optical flow is computed with a gradient–based technique. Fourth, optical flow vectors with small magnitudes are discarded. Fifth, vectors with similar locations, magnitudes, and directions are clustered together using a spatial consistency test. Sixth, similar optical flow vectors are extended temporally to make predictions about future optical flow locations, magnitudes, and directions in subsequent frames. The actual optical flow vectors that are consistent with those predictions are associated with a moving object. This algorithm was tested on images obtained with a videocamera mounted below the nose of a Boeing 737. The camera recorded two sequences containing a second flying aircraft. The algorithm detected the aircraft in 82% of the frames from the first sequence and 78% of the frames from the second sequence. In each sequence, the false alarm rate was zero. These results illustrate the effectiveness of using a comprehensive predictive technique when detecting moving objects.

Keywords: gradient, optical flow, object detection, computer vision

1 Introduction

The National Aeronautics and Space Administration (NASA) is currently collaborating with the commercial aircraft industry to develop a supersonic passenger airplane called the High Speed Civil Transport (HSCT). One issue being examined for HSCT is the replacement of the forward cockpit windows with synthetic displays. The imagery in these displays would be obtained from video cameras mounted outside the aircraft. A benefit of this configuration is that the video imagery can be examined with computer vision algorithms to determine if another aircraft is in the field of view.

The goal of this HSCT sub–project is to develop a computer vision algorithm to detect aircraft that are moving in the video images. One possible means of aircraft detection could be based on pattern recognition. However, because of the relatively small target size (fewer than 30 pixels at distances greater than 1.0 mile) and the background clutter, this technique was not chosen. Instead, optical flow (defined as the motion of brightness patterns in an image[1]) was selected as the basis for detection because the object is moving in the scene, thereby producing a different image velocity profile than the surroundings. The rationale of searching for variations in the optical flow pattern has proven useful for related types of object discrimination. For example, with images acquired with a moving camera, this principle has been used to detect static objects located at different distances than a stationary background[2-6]. In addition, this principle has been used for detecting moving objects against a stationary background[7-11]. The temporal characteristics of the optical flow profile can provide additional information about the reliability of the computations. A temporal approach (one that makes assumptions about the evolution of image velocity over a sequence of frames) has been used in the computation of accurate optical flow vectors in a static environment[12-14]. However, the purpose of those implementations was not to extract the locations of independently moving objects. Investigations that have used temporal information for this purpose have not fully exploited it. For example, a technique that separated different classes of motion did not include temporal predictions about the magnitude and direction of the optical flow vectors[15]. Instead, correspondence of features in temporally adjacent frames was based only on the *location* of the optical flow vectors. A second technique that separated different classes of motion did not fully utilize the temporal information because comparisons of optical flow vectors were made over only two sequential frames at a time[16]. Moreover, both of these approaches were based on images associated with a stationary camera, thereby restricting the ambiguities in optical flow that may occur with a moving camera[8].

In the current application, optical flow vectors are used to provide an estimate of the future locations, directions, and image speeds of a moving object (i.e., aircraft). This type of prediction technique validates the computation of future optical flow vectors, and is suited for conditions in which accurate optical flow computations may be difficult due to real, noisy images, such as those obtained with a moving camera directed at distant objects.

This paper first describes the algorithm, which is based on a gradient technique. Next, the paper describes an application of the algorithm for detecting a

target aircraft in video images. Finally, the results are quantified in terms of the hit rate and false alarm rate for different prediction parameters.

2 Description of the algorithm

The optical flow algorithm computes vectors associated with a translating object. To detect and track the target, six steps are required. These six steps are outlined in the block diagram in Fig. 1.

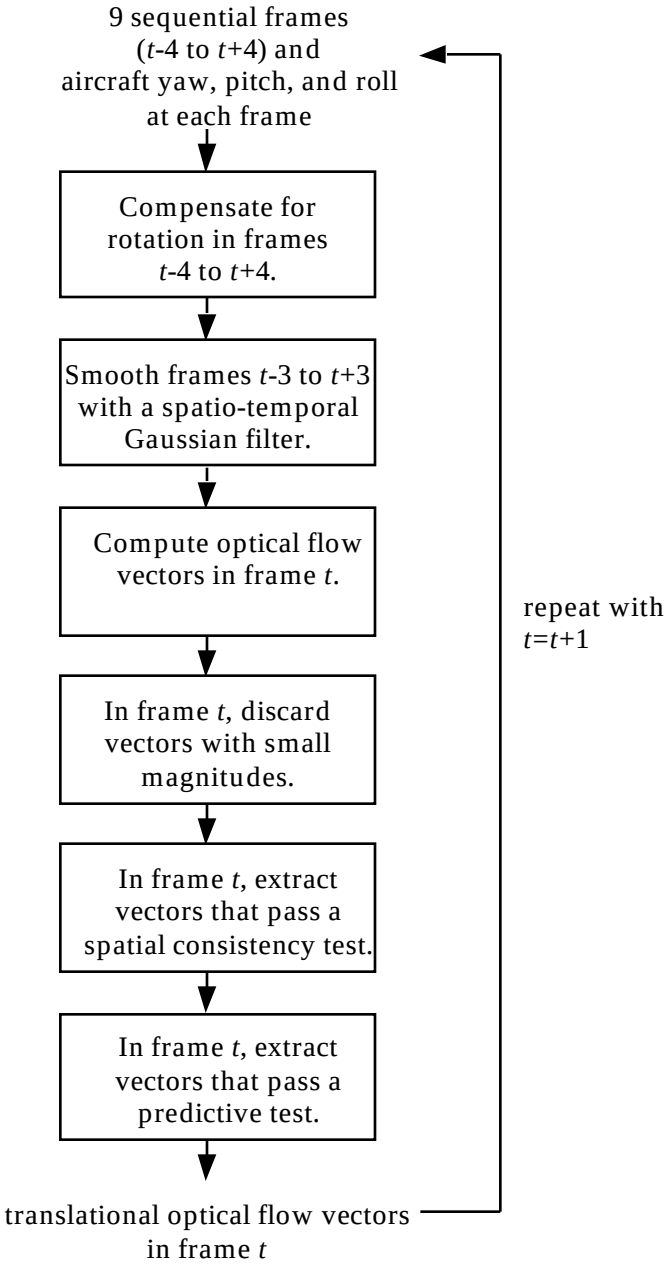

Fig. 1 *Block diagram of the steps in the predictive optical flow technique.*

In the first step, the algorithm compensates for image motion caused by camera rotation. Compensation is based on aircraft navigational parameters (pitch, roll, and yaw) recorded with the visual images themselves. Each pixel in an image surface is shifted according to the relative amount of aircraft rotation between it and the reference image frame. The reference image frame is simply the image frame at which the optical flow vectors are being computed. The eight frames prior to and subsequent to each reference frame are compensated at a time, so nine frames are used in each compensation process. (Nine frames are required because five frames are needed for the temporal support of the optical flow computations, and an additional four frames are needed in the temporal smoothing process.)

In addition to undergoing rotation effects, the images can also contain motion due to translation[17]. However, in this application image motion caused by aircraft translation is not compensated since the resulting optical flow vectors have such small magnitudes. The optical flow vectors caused by camera translation are negligible because the camera altitude was 4500 feet, the vertical field–of–view (FOV) was less than 10 deg, and the speed was 150 knots (about 173 mph), producing a maximum optical flow magnitude of 0.04 pixels / frame. (The optical flow vectors caused by rotation are not negligible because the rotation magnitudes are as large as 1.0 pixel / frame.)

In the second step, the image is smoothed with a spatiotemporal Gaussian filter with standard deviations of one pixel–frame. The spatial range is five pixels and the temporal range is five frames. These parameters are based on filter sizes used in previous implementations of the gradient–based approach[18].

In the third step, optical flow vectors are computed using the gradient constraint equation[1]:

$$\frac{\partial I}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\partial I}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}=-\frac{\partial I}{\partial t} \tag{1}$$

where $I = I(x,y,t)$ is the image intensity in gray levels (0-255), x and y are the spatial coordinates, and t is the frame time. The inherent assumption of Eq. (1) is that a given point in space produces constant brightness in the image as the point moves. Because this equation represents the gradient constraint at a given pixel (x,y) and frame (t), it can be rewritten in the following format:

$$I \cdot \mathbf{v} + I_t = 0 \tag{2}$$

where $I = [I_x \; I_y]$, I_x and I_y are the partial space–derivatives of I, I_t is the partial time–derivative of I, and $\mathbf{v}$ is the image velocity vector $[\mathrm{d}x/\mathrm{d}t \; \mathrm{d}y/\mathrm{d}t]^T$. In this equation, the partial derivatives can be estimated by using linear regression to compute the intensity change as a function of space and time surrounding each pixel (i.e., the partial derivatives). However, this process still leaves two unknown parameters (the optical flow components). Consequently, a second constraint is

required. The second constraint is based on the assumption that the velocity vector **v** is constant within a small region[19]. Using this assumption, Eq. (2) is applied within a small window to produce an over–constrained, linear system of equations[20]. From this constraint, the quadratic error function *E* associated with each optical flow vector is:

$$E(x,y,t) = \sum_{x'=-2}^{2}\sum_{y'=-2}^{2}\left[\, I(x+x',y+y',t)\cdot \mathbf{v}(x,y,t) + I_t(x+x',y+y',t)\right]^2 \qquad (3)$$

where $x', y' \in \{-2,...,2\}$ to account for the 5 x 5 window necessary for the second constraint. In addition, a weighting function is included to give more emphasis to the pixels in the center of the window than the periphery[21]. The resulting error equation is:

$$E(x,y,t) = \sum_{x'=-2}^{2}\sum_{y'=-2}^{2}\left[W^2(x',y')\right]\left[\, I(x+x',y+y',t)\cdot \mathbf{v}(x,y,t) + I_t(x+x',y+y',t)\right]^2 \qquad (4)$$

where

$$W^2(x',y') = w(x')w(y')$$

and

$$\begin{cases} w(-2) = w(2) = 0.0625 \\ w(-1) = w(1) = 0.25 \\ w(0) = 0.375 \end{cases} .$$

Consequently, each of the 25 image pixels in the window has a weight associated with it. The solution to Eq. (4) has the form:

$$\mathbf{A}^T\mathbf{W}^2\mathbf{A}\,\mathbf{v} = \mathbf{A}^T\mathbf{W}^2\mathbf{b} \qquad (5)$$

where

$$\mathbf{A} = \begin{bmatrix} I_x(x-2,y-2,t) & I_y(x-2,y-2,t) \\ I_x(x-2,y-1,t) & I_y(x-2,y-1,t) \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ I_x(x+2,y+1,t) & I_y(x+2,y+1,t) \\ I_x(x+2,y+2,t) & I_y(x+2,y+2,t) \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} W(-2,-2) & & & & 0 \\ & W(-2,-1) & & & \\ & & \ddots & & \\ & & & W(2,1) & \\ 0 & & & & W(2,2) \end{bmatrix}$$

$$\mathbf{b} = -\begin{bmatrix} I_t(x-2, y-2, t) \\ I_t(x-2, y-1, t) \\ \vdots \\ I_t(x+2, y+1, t) \\ I_t(x+2, y+2, t) \end{bmatrix}$$

As discussed by Barron, Fleet, and Beauchemin[18], the optical flow vector can then solved using:

$$\mathbf{v} = [\mathbf{A}^T\mathbf{W}^2\mathbf{A}]^{-1}\mathbf{A}^T\mathbf{W}^2\mathbf{b} \qquad (6)$$

This equation provides the optical flow vector at the pixel (x,y) at frame time t. Because of the spatial window used in the computation, optical flow vectors are not computed at points within two pixels of the frame border.

In the fourth step, optical flow vectors with small magnitudes are discarded. A threshold of 1.5 pixels / frame was chosen because this value corresponds with a target flying laterally to the camera at a relatively large distance and moderate speed, e.g., a target at 5 nautical miles (5.75 statute miles) with a speed of 300 knots (345 statute miles per hour). This threshold is small enough such that objects flying at closer distances or higher speeds would not be missed. Targets with smaller optical flow vectors (due to a greater distance or smaller speed) could be detected by reducing this threshold.

In the fifth step, spatially consistent optical flow vectors (called clustered vectors) are extracted from the optical flow patterns using a spatial consistency test[22-24]. Vectors that fail the spatial consistency test are discarded. Each of the nine vectors in a 3 x 3 window is defined as a clustered vector if the eight boundary vectors have similar magnitudes and directions as the center vector. To pass the spatial consistency test, the following equations are applied to every pixel:

$$|M(i) - M_c| < M_t \qquad \text{for } i=1,...,8 \text{ in a window} \qquad (7)$$

$$|\theta(i) - \theta_c| < \theta_t \qquad \text{for } i=1,...,8 \text{ in a window} \qquad (8)$$

where $M(i)$ and $\Theta(i)$ are the magnitude and direction of the optical flow vector i, M_c and Θ_c are the magnitude and direction of the optical flow vector in the window center, and M_t and Θ_t are the magnitude and direction thresholds (0.2 pixels/frame and 10 deg, chosen empirically). Reducing the threshold magnitudes can reduce the false alarm rate. However, the more robust technique used in this project reduces false alarms with a predictive technique.

In the sixth step, the predictive technique is used to estimate the future locations, magnitudes, and directions of the spatially consistent optical flow vectors. Each of the clustered vectors in a given frame is extended temporally to make predictions about the location, magnitude, and direction of optical flow vectors in the next P frames:

$$X'(t+\Delta t) = X(t) + \mathbf{v}(t)\Delta t \tag{9}$$

where $X'(t+\Delta t)$ is the predicted location of the optical flow vector in frame $t+\Delta t$ and $X(t)$ is the computed location of the optical flow vector in the current frame t. The magnitude and direction of a predicted vector at $t+\Delta t$ are the same as the those of a reference vector at t. For example, if frame 15 contains a clustered optical flow vector at (x,y) coordinate (10,10) with a magnitude of 1.0 pixel/frame and a direction of 90 deg, then a prediction is made that frame 16 will contain an optical flow vector at (x,y) coordinate (10,11) with the same magnitude and direction. If P=5, predictions are also made about frames 17, 18, 19 and 20. Because of this technique, a given frame contained predictions from the previous P frames. For example, if P=5, frame 16 not only contains predictions based on frame 15, it also contains predictions based on frames 11, 12, 13, and 14. Therefore, in each frame two pieces of information are available: 1) the predicted optical flow vectors based on the previous P frames and 2) the actual optical flow vectors based on steps 1 through 5.

In the predictive step, if the clustered optical flow vectors in a given frame are temporally consistent with the predicted optical flow vectors from the previous frames, the clustered vectors are marked as translational optical flow vectors. (For this application, translational optical flow vectors are defined as vectors associated with an independently moving object with a velocity component perpendicular to the camera line of sight.) The temporal consistency test is defined as follows:

1) The location of acceptable translational optical flow vectors is restricted to a 5x5 window surrounding the predicted pixel location.
2) The magnitude is restricted to within ±0.25 pixels/frame of the predicted magnitude.
3) The direction is restricted to within ±10 deg of the predicted direction.

To be accepted as a valid translational optical flow vector, all three components of this test are required to be satisfied for P' of the previous P frames. As presented in the next section, the hit rate and false alarm rate can be modified by adjusting these two parameters (P' and P).

A simplified example of the predictive step is shown in Fig. 2. In this example, four sequential frames are shown, each of which has a clustered optical flow vector in it. From each frame, predictions are made for the next three frames

(so *P*=3). In frame 4, the clustered optical flow vector is temporally consistent with two of the previous three frames. If P' is defined as less than three in this example, the clustered optical flow vector in frame 4 passes the temporal consistency test, and is therefore defined as a translational optical flow vector.

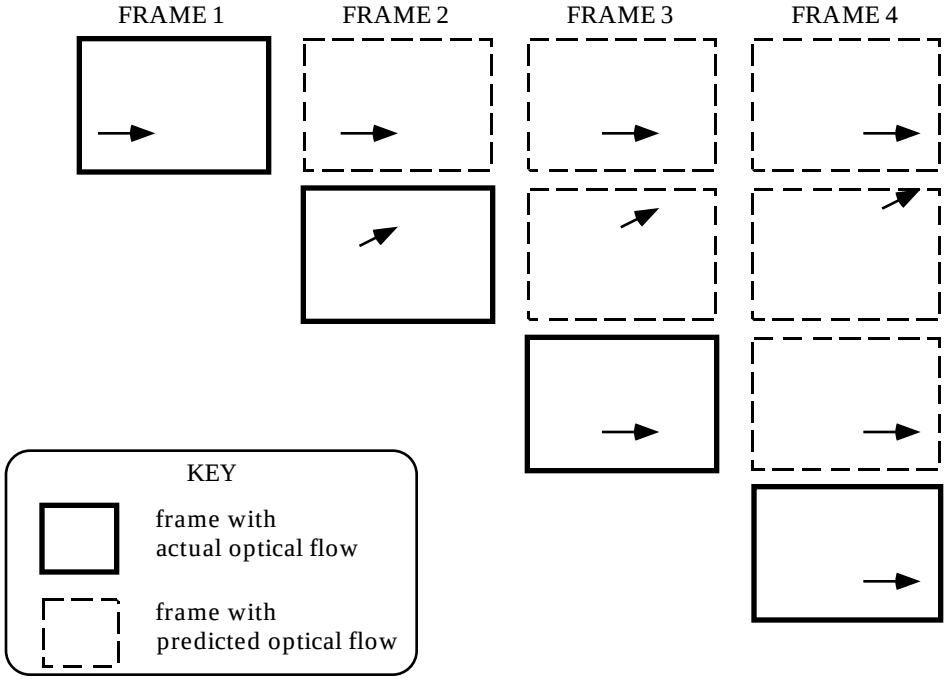

Fig. 2 *Simplified schematic illustrating the predictive portion of the algorithm. In this example, a clustered optical flow vector is present in four sequential frames. In each frame, predictions are made about the location, magnitude, and direction of the optical flow vector in the next three frames (so P=3). (For simplicity, predictions are not shown beyond frame 4.) In frame 4, the actual clustered optical flow vector is consistent with the predicted optical flow vectors from two of the previous three frames. Consequently, if P' is defined as 0, 1, or 2 in this example, the clustered vector in frame 4 would pass the temporal consistency test. If P' is defined as 3, it would fail the temporal consistency test.*

3 Implementation of the algorithm

To collect test data for the algorithm, flight tests were conducted at NASA Langley Research Center on April 30 and May 23, 1997. In these tests, a Kodak Megaplus ES 1.0 monochrome camera was mounted below the nose of a Boeing 737 to record a Beech King Air 200 flying in different trajectories. The Boeing 737 flew at a 150 knots and the Beech King Air flew at 200 knots. The ground was continuously within the FOV, providing a source of background clutter. The camera recorded images at a rate of 30 frames per second in 8–bit format, producing 256 gray levels. The camera images were recorded in S–VHS videotape format and then digitized on a Silicon Graphics (SGI) Onyx computer. The final images had a 13 x 9.75 deg FOV

with a pixel array of 646 x 486 (resulting in 50 pixels / deg). A time code was inserted on the video track to monitor each individual frame.

In addition to video imagery, navigational data were recorded. These data were used to compensate for host aircraft rotation, as discussed in the previous section. The navigational data was acquired at a different rate (20 Hz) than the images (30 Hz), so interpolation was used to estimate the pitch, roll, and yaw at each frame time. The rotation angles of the camera are independent of the point on the aircraft at which they were measured (because the camera was fixed to the rigid body aircraft). Camera translation caused by the rotation was negligible due to the altitude of the aircraft.

A typical image recorded during the flight tests is shown in Fig. 3. In this condition, the target aircraft was located at a distance of about 1.2 nautical miles (nmi) from the camera on the Boeing 737.

Fig. 3 *Typical image recorded during the flight tests. In this image, the target is located in the lower portion of the image about 1 cm above the number 15.*

In Fig. 3, the target is located against the earth in the lower portion of the image. In this image, the signal–to–noise ratio (SNR) is 28 dB, which is typical for an image in this sequence. The signal (S) is computed from the difference between the target intensity values and the mean background intensity (B)[25-26]. The target has dimensions of 9 x 3 pixels (a total of 27 pixels). The background is defined as the border (four pixels wide) surrounding the target, resulting in background dimensions of 17 x 11 pixels. The SNR is defined as follows:

$$\mathrm{SNR} = 20\log_{10}\frac{\sqrt{\sum_{i=1}^{T}\left(I_i - B\right)^2}}{\sigma_B} \tag{10}$$

where I_i is the image intensity at target pixel i, T is the total number of target pixels (27), B is the mean background intensity, and σ_B is the standard deviation of the background intensity.

After camera motion compensation (step 1) and image smoothing (step 2), optical flow vectors are computed with the gradient–based technique (step 3). Next, vectors with small magnitudes are discarded (step 4), and a spatial consistency test is applied to isolate the vectors with characteristics corresponding to a translating target (step 5). In general, the spatial consistency test did not remove all false alarms, as shown in Fig. 4.

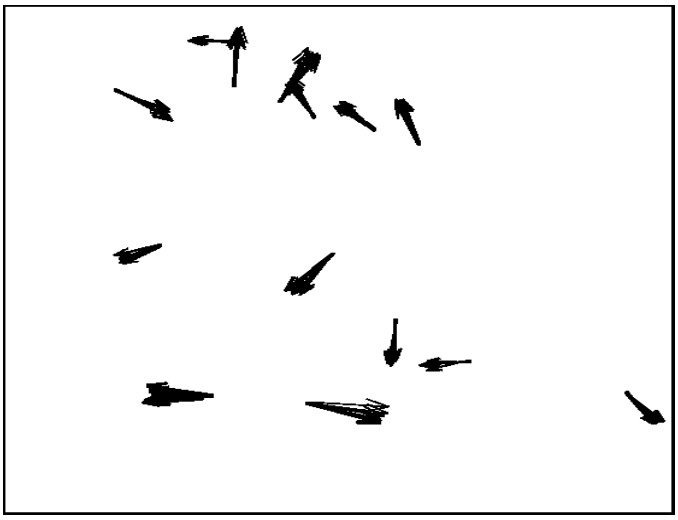

Fig. 4 *Clustered optical flow remaining after the spatial consistency test. The 162 optical flow vectors shown here (in 14 clusters) correspond to the image in Fig. 2. These vectors are scaled by a factor of 25 for clarity.*

In the final step, the predictive technique is used to reduce the spurious clustered optical flow vectors that are computed with the spatial consistency test. An example of the predictive results is shown in Fig. 5 for a typical frame (frame #201, which corresponds to Fig. 2). In this example, predictions are made over twelve frames, so P=12. Consequently, the optical flow in frame #201 is predicted from the clustered optical flow vectors in the previous 12 frames (frames #189-200). The predictive test is based on correspondence for three of the twelve frames, so P'=3. In Fig. 5, the white areas represent regions in which the predicted optical flow vectors from at least 3 of the previous 12 frames are in spatial correspondence.

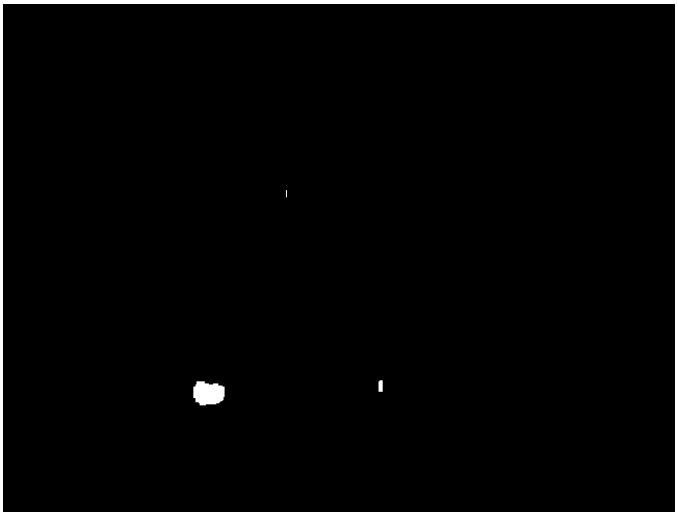

Fig. 5 *Predicted positions of the translational flow vectors in the frame corresponding to Fig. 2. In 3 of the previous 12 frames, translational vectors are predicted to occur at the white regions.*

If the clustered optical flow vectors in frame #201 are located in the white regions of Fig. 5, they pass the location component of the temporal consistency test. If the clustered optical flow vectors also pass the magnitude and direction components of the temporal consistency test, they are defined as translational optical flow vectors associated with an independently moving object. Those vectors are shown in Fig. 6.

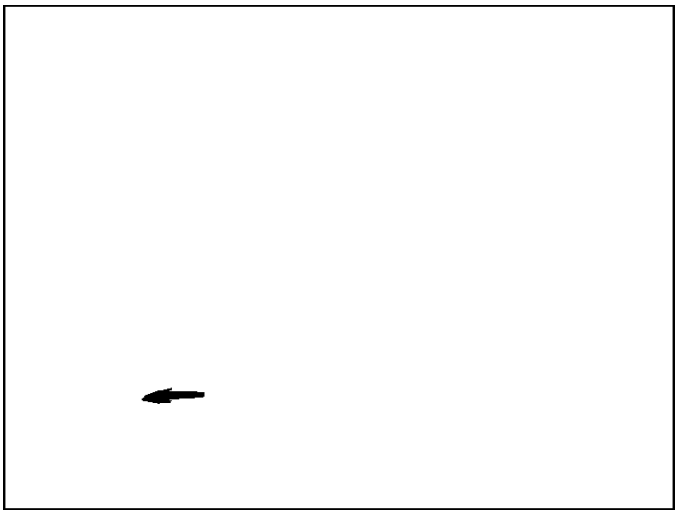

Fig. 6 *Translational optical flow computed with the predictive technique for the frame corresponding to Fig. 2. These 27 translational flow vectors represent the clustered flow vectors with locations, magnitudes, and directions within the acceptable range dictated by the predictive test. These vectors are scaled by a factor of 25 for clarity.*

Figs. 4–6 illustrate typical results for a single frame in a nine second (270 frame) sequence containing a moving aircraft. The results over all frames are summarized in Table 1 using a receiver operating characteristic (ROC) analysis[27]. ROC analysis is based on the number of frames (s) that contain a region with the target and the number of frames (n) that contain a region without the target. The target is located in 217 frames (frames 34-250), so s=217. Every image contains regions without the target, so n=270. The term S refers to the number of frames with hits and N refers to the number of frames without hits.

		DETECTION: yes	DETECTION: no
TARGET	yes	HIT p(S | s) *82%*	MISS p(N | s) *18%*
	no	FALSE ALARM p(S | n) *0%*	CORRECT REJECTION p(N | n) *100%*

Table 1 *Stimulus–response matrix for classification of the algorithm performance. As an example, probabilities for P=12 and P'=3 are shown in the lower right corner of each box.*

An image is defined as a hit if at least one of the detected target pixels corresponds with an actual target pixel. If no hits occur, the image is classified as a miss. Consequently, p(S | s) + p(N | s) = 1.0. An image is defined as a false alarm if the algorithm marks at least one non–target pixel as a hit. Otherwise the image is labeled a correct rejection. Consequently, p(S | n) + p(N | n) = 1.0.

Table 2 shows the results for several durations (P) in the predictive technique. For each duration, the required overlap (P'/P) of location, magnitude, and direction ranges from 0% to 100%. The hit rate and false alarm rate both decrease with increased overlap (P'/P) and increased duration (P). The results indicate that a 25% overlap with a duration of P=12 frames provide the highest hit rate (82%) with zero false alarms. When predictive techniques are not used, all optical flow vectors that pass the spatial consistency test are used to mark the target. In that case, the hit rate is 97%, meaning that the target is correctly detected in 97% of the frames that contained a target. However, the false alarm rate is 100%, meaning that in every frame at least one translational optical flow vector does not correspond to the target location.

	FRAMES *P*			
P'/P (%)	0 (off)	4	8	12
0	97/100	---	---	---
25	---	86/34	85/1	82/0
50	---	76/1	73/0	69/0
75	---	54/0	33/0	18/0
100	---	23/0	4/0	1/0

Table 2 *ROC results ([hits]/[false alarms]) for the first scenario. These results demonstrate the effect of different number of frames used in the predictive technique. The term P refers to the number of frames used in the prediction process. The term P' refers to the number of those frames which satisfy the temporal consistency test. For example, if P'/P = 75% and P=8 frames, then 6 of those 8 frames must be satisfied by the temporal consistency test. If the prediction technique is not used (P=0) false alarms occur in every frame.*

As an additional test of the algorithm, the results were examined for a second sequence (containing 160 frames) containing a moving aircraft. This sequence is similar to the first except that the background of the target aircraft is the sky instead of the earth. In this case, the typical SNR is 25dB, which is lower than the SNR in the first sequence. The hit rate and false alarm again decrease with increased overlap (*P'*/*P*) and increased duration (*P*). Once again, a 25% overlap with a duration of *P*=12 frames provides the highest hit rate (78%) for zero false alarms. The results for both predictive parameters (*P'* and *P*) are shown in Table 3.

	FRAMES *P*			
P'/P (%)	0 (off)	4	8	12
0	96/100	---	---	---
25	---	96/14	93/1	78/0
50	---	77/1	35/0	11/0
75	---	28/0	3/0	0/0
100	---	3/0	0/0	0/0

Table 3 *ROC results ([hits]/[false alarms]) for the second scenario. The parameters here are identical to those in Table 2.*

The results from these two sequences illustrate the effectiveness of the predictive technique, which reduces the false alarms that remain after the spatial consistency test (step 5). An alternative means of reducing the false alarms while maintaining a high hit rate is to modify the spatial consistency parameters. For

example, if the magnitudes and orientation thresholds in Eqs. (7) and (8) are reduced by 25%, the hit rate from the spatial consistency test in the first scenario is 82% (the same hit rate that can be obtained with the predictive test). However, the false alarm rate is still high (98%), whereas the false alarm rate from the predictive technique is 0%. Using the same principle for the spatial consistency test in the second scenario, the hit rate becomes 78% (the same hit rate that can be obtained with the predictive test). But once again, the false alarm rate is high (60%). These results illustrate that the predictive technique provides better performance than a purely spatial consistency technique.

4 Discussion

The predictive technique provides an effective means of reducing false alarms in optical flow computations. Not surprisingly, both the hit rate and false alarm rate decrease as the required overlap (P'/P) of the predictive technique increases. This occurs because the temporal consistency test adds an additional level of stringency beyond the spatial consistency test in the computation of translational optical flow vectors. Temporal aliasing and quantization may limit the effectiveness of the optical flow computations[18], thereby producing inaccurate optical flow vectors in the region of the target. Consequently, the stringent predictive test provides a more accurate assessment of those vectors by discarding those that are not temporally consistent. The benefit of this approach is that all of the false alarms can be discarded. For both sequences, the highest hit rate with a zero false alarms occur when three of the previous 12 frames are in correspondence with the current frame. These parameters provide a strong hit rate while eliminating the false alarms.

Another possible mechanism for reducing false alarms would be pre–processing to ignore regions with low contrast. However, a low contrast region could still correspond to the target because of hazy weather conditions. In addition, the size, projected shape, and surface reflection properties of potential aircraft are unknown, thereby making contrast assumptions inappropriate. Consequently, optical flow computations are not restricted to high contrast regions in the image.

This algorithm was designed to produce accurate results without regards to the time necessary to detect the target. Currently, the algorithm is being refined to work in real–time with a minimal decrement in performance. For example, the weights associated with the gradient–constraint equation can be changed to byte format (e.g., 1/8, 4/8, and 5/8), thereby minimizing the number of bits necessary in the hardware implementation.

A CCD camera with 256 gray levels and a 30 Hz provides an acceptable sensor for this application. In the current example with this camera, the target subtended about 27 pixels (9 x 3). For greater target distances, the camera FOV can be reduced so that the same number of pixels are driven by target. However, the trade–off is that high frequency vibration can cause an unacceptable amount of motion in the image surface, making spatio–temporal smoothing less effective. Consequently, the

accurate measurement of aircraft vibration and motion should be considered when decreasing the camera FOV to provide more target pixels.

The predictive technique described here is similar to the temporal coherence theory proposed by Grzywacz, Watamaniuk, and McKee[28] to explain human motion perception. This theory accounted for psychophysical experiments in which subjects attempted to detect a single dot that moved with consistent motion while surrounding dots moved with random motion. Speed information was discarded in their implementation. The predictive technique is different in that speed information is used in validating the computation of optical flow vectors. Another difference is in terms of the velocity signals themselves. With biological temporal coherence theory, smoothing was applied to the outputs of directionally tuned cells for biological plausibility, resulting in additional noise to the system. With the predictive technique, the velocity signals are used directly in coherence estimation.

A related approach to temporal coherence was taken by Kruger, Enkelmann, and Rossle[22] for detecting automobiles on the road. Their key assumptions were that camera rotation was zero and the ground surface was planar. Kalman filtering was used to make depth estimates and track optical flow vectors. Tracking of potential obstacles was terminated if there were no appropriate optical flow measurements. Because optical flow measurements are inherently noisy, the predictive technique used in the current algorithm does not require accurate optical flow measurements in every frame. Consequently, this technique provides a more robust means for detection of moving objects.

Acknowledgments

The author wishes to thank Albert Ahumada, Mary Kaiser, and Lee Stone for their helpful comments. This project was supported by a National Research Council associateship award to the author and by NASA RTOP# 537-08-22.

References

[1] B.K.P. Horn and B.G. Schunck, "Determining optical flow," *Art. Intell.* **17**, 185-203 (1981).

[2] T. Camus, D. Coombs, M. Herman, and T. Hong, "Real–time single–workstation obstacle avoidance using only wide–field flow divergence," *Int. Conf. on Patt. Recog. Appl. and Robotic Sys.* 323-330 (1996).

[3] A. Dev, B.J.A. Krose, L. Dorst, and F.C.A. Groen, "Observer curve and object detection from optic flow," *SPIE Intelligent Robots and Computer Vision XIII: 3D Vision, Product Inspection, and Active Vision*, 38-49 (1994).

[4] W. Enkelmann, "Obstacle detection by evaluation of optical flow fields from image sequences," *Image and Vision Comp.* **9**, 160-168 (1991).

[5] M. Ilic, S. Masciangelo, and E. Pianigiani, "Ground plane obstacle detection from optical flow anomalies: a robust and efficient implementation," *Proc. Intelligent Vehicles '94 Symposium*, 333-338 (1994).

[6] R.C. Nelson and J. Aloimonos, "Obstacle detection using flow field divergence," *IEEE Trans. Patt. Analy. and Machine Intell.* **11**, 1102-1106 (1989).

[7] Y. Barniv, "Dynamic programming solution for detecting dim moving targets," *IEEE Trans. Aerospace and Elect. Sys.* **21** 144-156 (1985).

[8] D.J. Heeger and G. Hager, "Egomotion and the stabilized world," *IEEE Second Int. Conf. Computer Vision*, 435-440 (1988).

[9] R.C. Nelson, "Qualitative detection of motion by a moving observer," *IEEE Conf. on Computer Vision and Patt. Recog.* 173-178 (1991).

[10] W.B. Thompson, P. Lechleider, and E.R. Stuck, "Detecting moving objects using the rigidity constraint," *IEEE Trans. Patt. Analy. Machine Intell.* **15**, 162-166 (1993).

[11] Z. Zhang, O.D. Faugeras, and N. Ayache, "Analysis of a sequence of stereo scenes containing multiple moving objects using rigidity constraints," *IEEE Second International Conf. Computer Vision*, 177-186 (1988).

[12] M.J. Black and P. Anandan, "Robust dynamic motion estimation over time," *Proc. Computer Vision and Pattern Recognition*. 296-302 (1991).

[13] M.J. Black, "Recursive non–linear estimation of discontinuous flow fields," *Proc. 3rd European Conf. Computer Vision*, 138-145 (1994).

[14] T.M. Chin, M.R. Luettgen, W.C. Karl, A.S. Willsky, "An estimation theoretic perspective on image processing and the calculation of optical flow," in *Motion Analysis and Image Sequence Processing*, M.I. Sezan and R.L. Lagendijk, Eds., pp. 23-51, Kluwer Academic Publishers, Boston (1993).

[15] H. Chen and Y. Shirai, "Detecting multiple image motions by exploiting temporal coherence of apparent motion," *IEEE Comp. Vis. Patt. Recog.* 899-902 (1994).

[16] Y. Mae, Y. Shirai, J. Miura, and Y. Kuno, "Object tracking in cluttered backgrounds based on optical flow and edges," *Proc. 13th Int. Conf. Pattern Recognition*, **1** 196-200 (1996).

[17] H.C. Longuet–Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proc. Royal Soc. Lond. B* **208**, 385-397 (1980).

[18] J.L. Barron, D.J. Fleet, and S.S. Beauchemin, "Performance of optical flow techniques," *Inter. J. Comp. Vis.* **12**, 43-77 (1994).

[19] C.L. Fennema and W.B. Thompson, "Velocity determination in scenes containing several moving objects," *Comp. Graphics Image Proc.* **9**, 301-315 (1979).

[20] B.D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. 7th Inter. Joint Conf. Art. Intell.* **2**, 674-679 (1981).

[21] E.P. Simoncelli, E.H. Adelson, and D.J. Heeger, "Probability distributions of optical flow," *IEEE Conf. Comp. Vis. Patt. Recog.* 310-315 (1991).

[22] W. Kruger, W. Enkelmann, S. Rossle, "Real–time estimation and tracking of optical flow vectors for obstacle detection," *Proc. Intelligent Vehicles '95 Symposium*, 304-309 (1995).

[23] J.W. McCandless and M.K. Kaiser, "Obstacle detection in air–to–air images," *SPIE Enhanced Synthetic Vis.* **3088**, 40-49 (1997).

[24] Y. Tang and R. Kasturi, "Tracking moving objects during low altitude flight," *Machine Vision and App.* **9**, 20-31 (1996).
[25] E.R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*, Academic Press, San Diego (1997).
[26] B.K.P Horn, *Robot Vision*, The MIT Press, Cambridge, MA (1986).
[27] D.M. Green and J.A. Swets, *Signal Detection Theory and Psychophysics*, Robert E. Krieger Publishing Co., Huntington (1966).
[28] N.M Grzywacz, S.N.J. Watamaniuk, S.P. McKee, "Temporal coherence theory for the detection and measurement of visual motion," *Vis. Res.* **15**, 3181-3203 (1995).